Length Frequency

Alex J. Benecke

December 19, 2017

Data Preparation

I will now compare the length frequency distribution for largemouth bass obtained in the nearshore electrofishing survey during 2013 - 2016.

```
lmb <- read.csv("Data/Clean-Data/2012-2016_nearshore-survey-largemouth-bass_CLEAN.csv") %>%
    arrange(Year, FID, Length)
lmb$fyr <- as.factor(lmb$fyr)

str(lmb)
```

```
## 'data.frame':    496 obs. of  16 variables:
##  $ Year  : int  2012 2012 2012 2012 2012 2012 2012 2012 2012 2012 ...
##  $ Site  : int  18 18 18 18 18 18 18 18 18 18 ...
##  $ FID   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ Weight: num  8 10 10 30 25 20 40 155 145 170 ...
##  $ Length: int  72 82 85 108 110 115 119 220 220 230 ...
##  $ AC    : int  2 2 2 2 2 2 2 2 3 3 ...
##  $ AGE   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ SexCon: int  NA NA NA NA NA NA NA NA NA NA ...
##  $ Sex   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ Delts : logi  NA NA NA NA NA NA ...
##  $ logW  : num  0.903 1 1 1.477 1.398 ...
##  $ logL  : num  1.86 1.91 1.93 2.03 2.04 ...
##  $ fyr   : Factor w/ 5 levels "2012","2013",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ Ws    : num  3.56 5.44 6.12 13.41 14.24 ...
##  $ Wr    : num  225 184 163 224 176 ...
##  $ gcat  : Factor w/ 4 levels "preferred","quality",..: 4 4 4 4 4 4 4 3 3 3 ...
```

```
headtail(lmb)
```

```
##     Year  Site FID Weight Length AC AGE SexCon Sex Delts     logW     logL
## 1   2012    18  NA      8     72  2  NA     NA  NA    NA 0.903090 1.857332
## 2   2012    18  NA     10     82  2  NA     NA  NA    NA 1.000000 1.913814
## 3   2012    18  NA     10     85  2  NA     NA  NA    NA 1.000000 1.929419
## 494 2016    15 130    305    266  3   2      8   2    NA 2.484300 2.424882
## 495 2016    15 131    282    261  3   2      3   1    NA 2.450249 2.416641
## 496 2016 15972 132    971    395  3   7      3   1    NA 2.987219 2.596597
##      fyr         Ws       Wr      gcat
## 1   2012   3.556717 224.9266  substock
## 2   2012   5.443933 183.6907  substock
## 3   2012   6.123337 163.3097  substock
## 494 2016 256.234496 119.0316     stock
## 495 2016 240.804379 117.1075     stock
## 496 2016 934.678588 103.8860 preferred
```

```
unique(lmb$Year)  ### See that there is no 2013
```

```
## [1] 2012 2013 2014 2015 2016
```

Lets create a new variable for 20 mm length bins.

```
lmb %<>% mutate(lcat20 = lencat(Length, w = 20))
headtail(lmb)
```

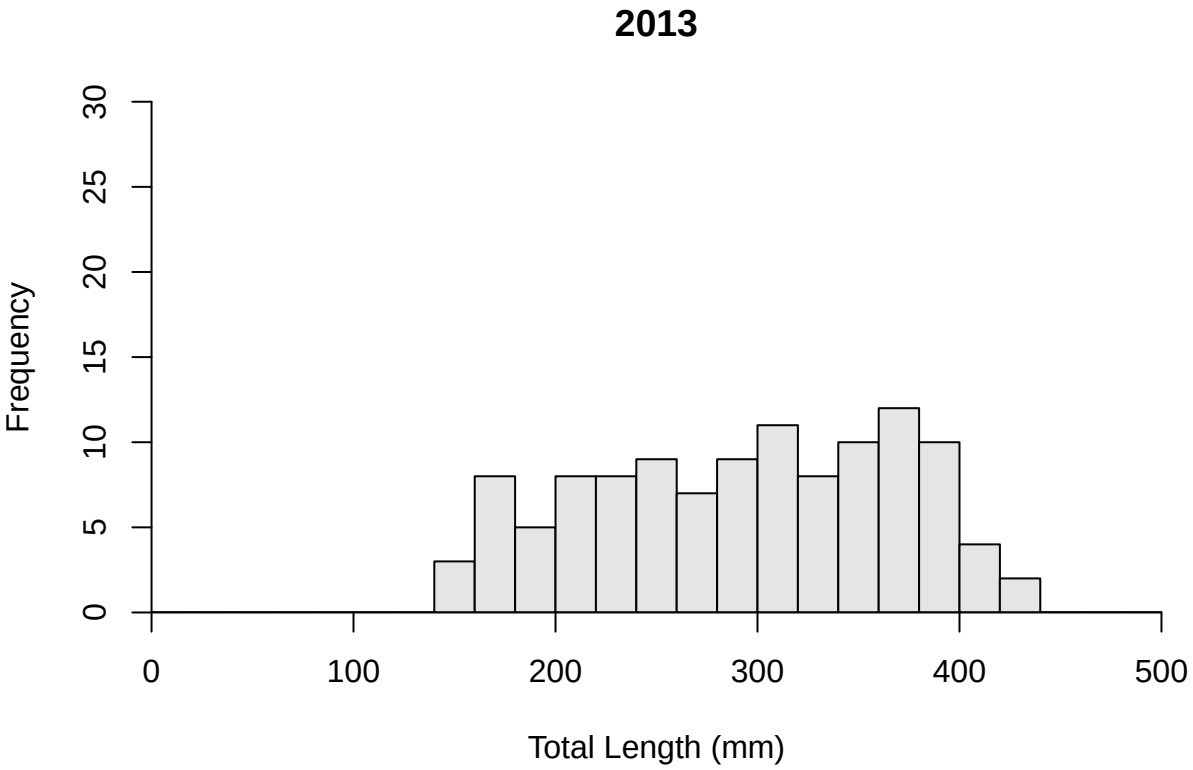
```
##      Year  Site FID Weight Length AC AGE SexCon Sex Delts     logW     logL
## 1    2012    18  NA      8     72  2  NA     NA  NA    NA 0.903090 1.857332
## 2    2012    18  NA     10     82  2  NA     NA  NA    NA 1.000000 1.913814
## 3    2012    18  NA     10     85  2  NA     NA  NA    NA 1.000000 1.929419
## 494 2016    15 130    305    266  3   2      8   2    NA 2.484300 2.424882
## 495 2016    15 131    282    261  3   2      3   1    NA 2.450249 2.416641
## 496 2016 15972 132    971    395  3   7      3   1    NA 2.987219 2.596597
##       fyr         Ws       Wr      gcat lcat20
## 1    2012   3.556717 224.9266  substock     60
## 2    2012   5.443933 183.6907  substock     80
## 3    2012   6.123337 163.3097  substock     80
## 494 2016 256.234496 119.0316     stock    260
## 495 2016 240.804379 117.1075     stock    260
## 496 2016 934.678588 103.8860 preferred    380
```

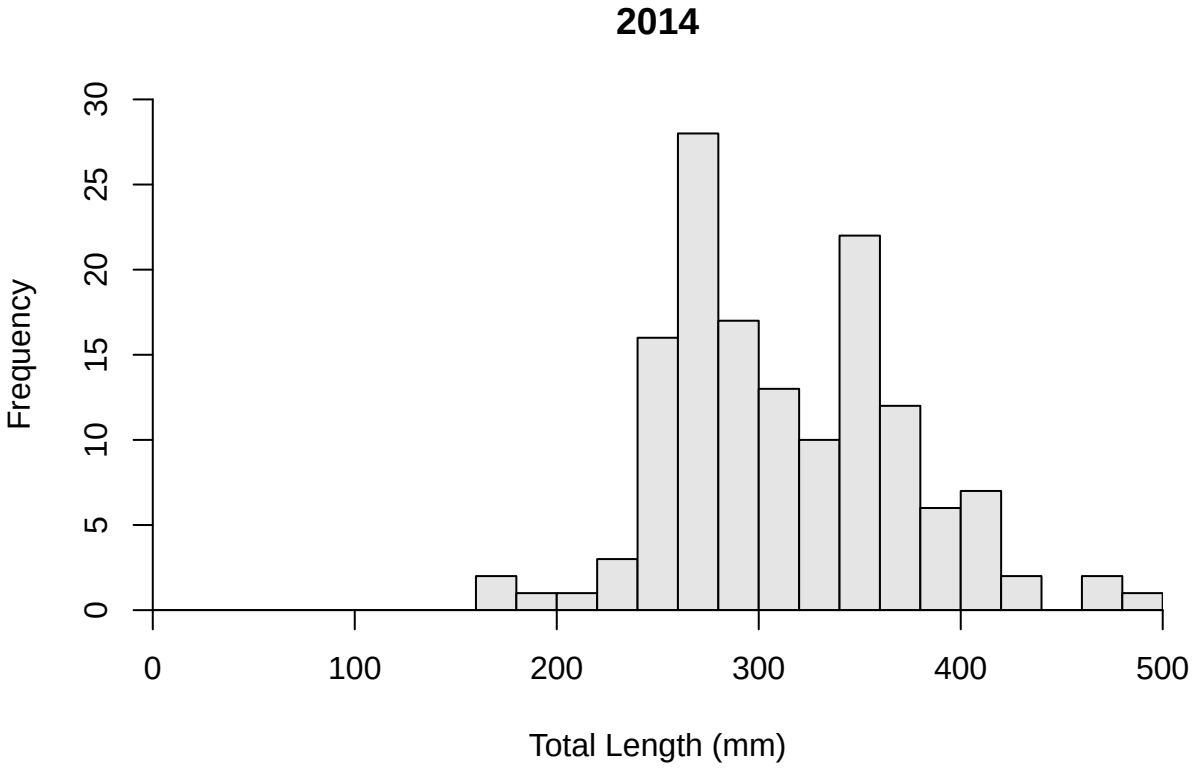
Now I want to separate out the years. I will throw out the year 2012 because samples from this years were not collected using the same procedures as in subsequent years. On ly large LMB from 2012 had length weigh data.

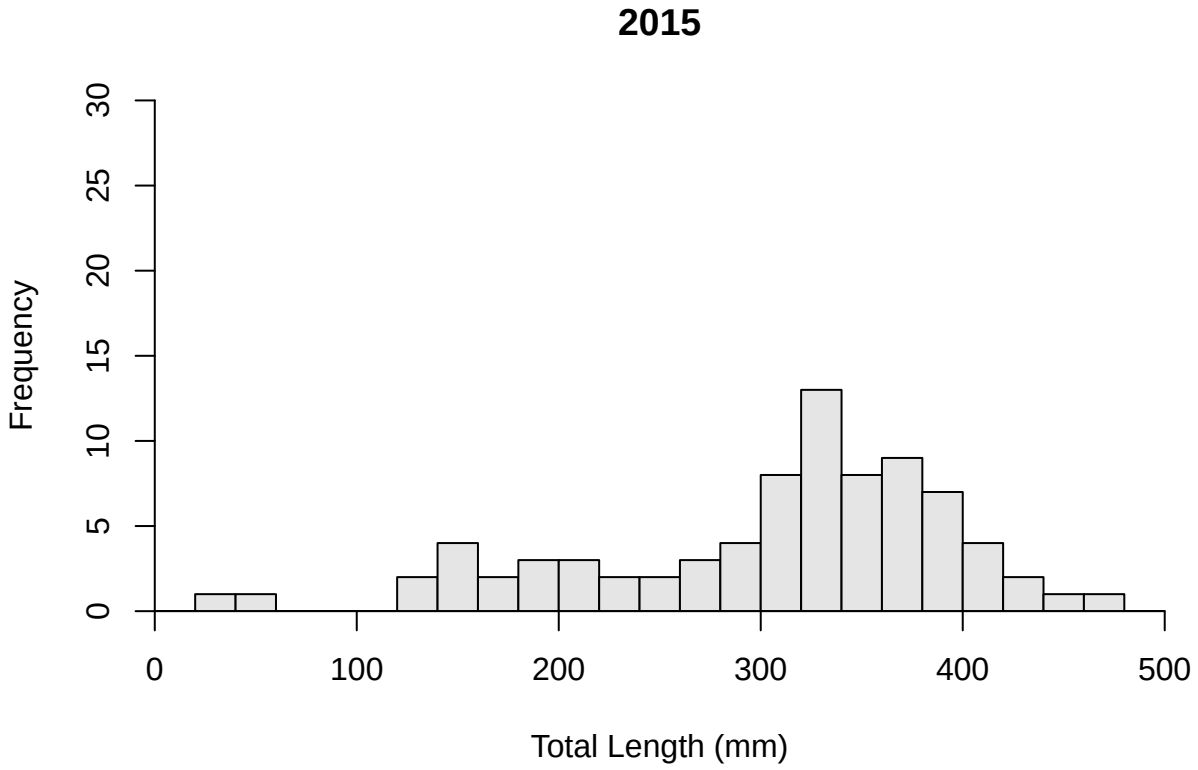
```
lmb.12 <- filter(lmb, Year == 2012)
lmb.13 <- filter(lmb, Year == 2013)
lmb.14 <- filter(lmb, Year == 2014)
lmb.15 <- filter(lmb, Year == 2015)
lmb.16 <- filter(lmb, Year == 2016)
```

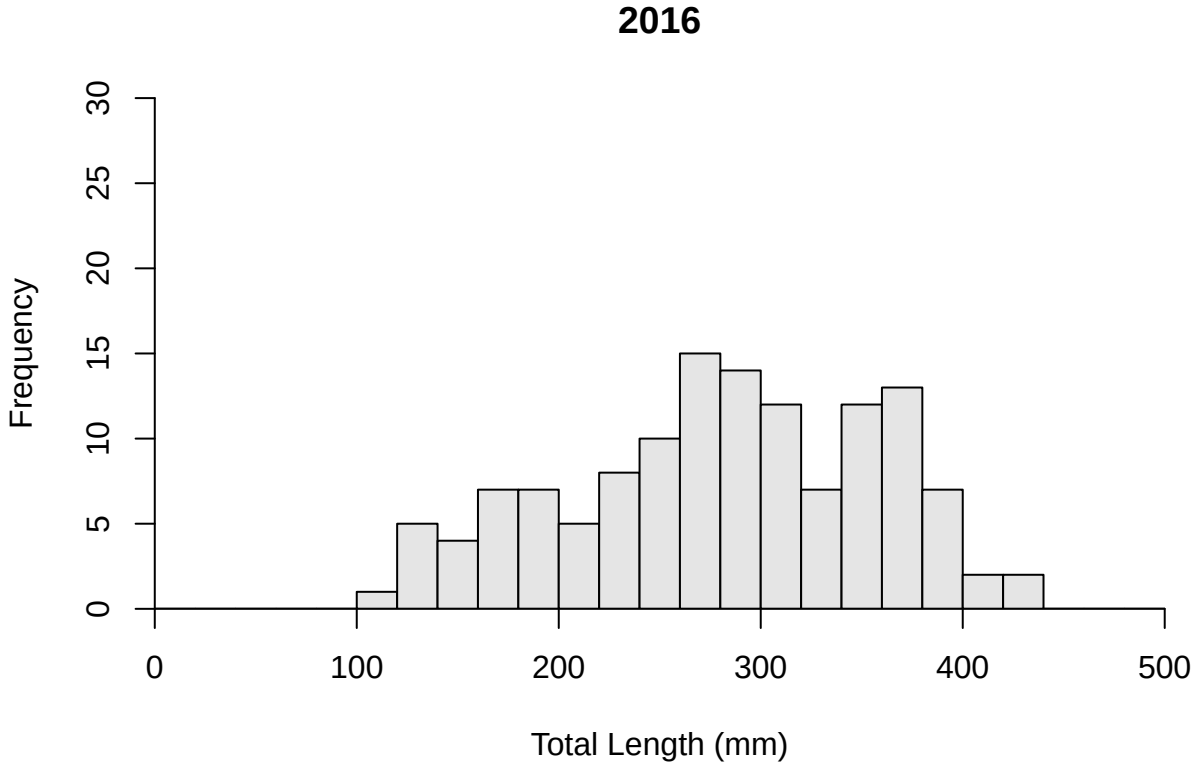
Length Frequency Distribution

Lets view a quick histogram of the frequency of fish in each length bin.





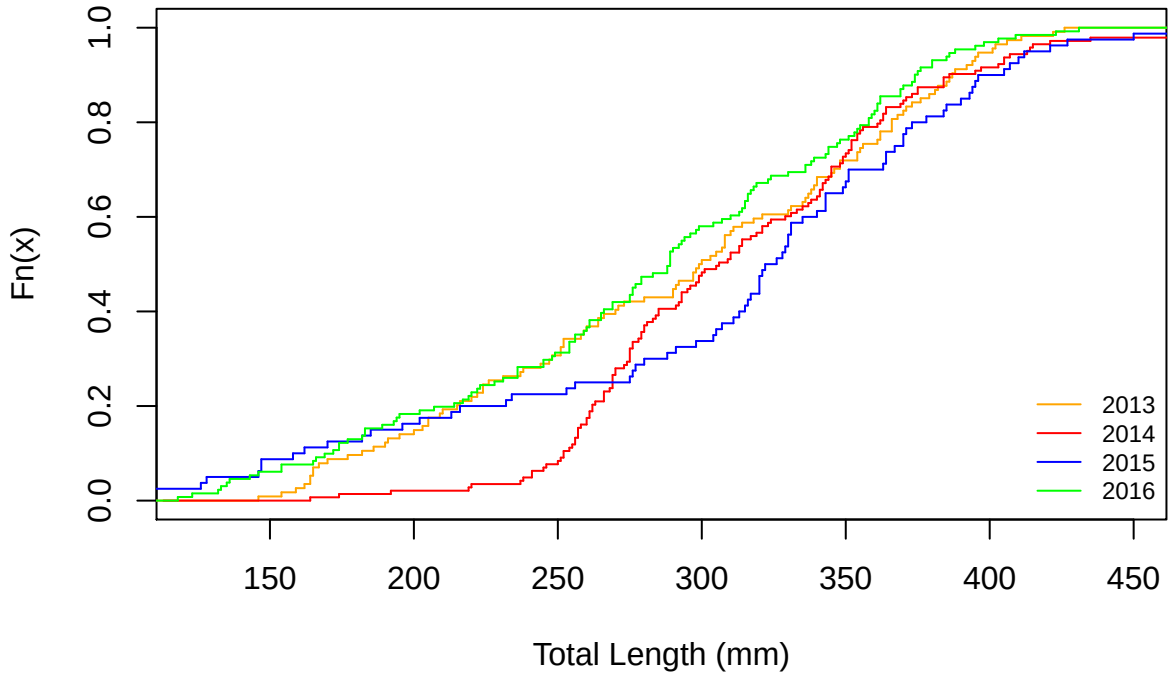




There may be a problem where small fish (<100 mm) are not being captured by our gear dispite some being present in 2015. 2014 and 2016 look fairly similar to me and 2015 doesn't look too far off. With my limited expirience in these matters I would have to say the largemouth bass population looks stable. But lets continue to check this in a less qualitative manner.

Cumulative Frequencies

Lets look at the empirical cumulative distribution function (ECDF). This is the proportion of fish less than each observed length. This should help me compare the length frequency distributions between years.



Compare Length Frequency Between Years

Kolmogorov-Smirnov Test

```
(D <- c(ks.13.14$statistic[[1]], ks.13.15$statistic[[1]], ks.13.16$statistic[[1]],
    ks.14.15$statistic[[1]], ks.14.16$statistic[[1]], ks.15.16$statistic[[1]]))
```

```
## [1] 0.23721016 0.19517544 0.09997322 0.19003497 0.24747771 0.24265267
```

```
(yrs <- c("13-14", "13-15", "13-16", "14-15", "14-16", "15-16"))
```

```
## [1] "13-14" "13-15" "13-16" "14-15" "14-16" "15-16"
```

```
(p.yr <- data.frame(yrs, D, p.val))
```

```
##     yrs          D       p.val
## 1 13-14 0.23721016 0.007939765
## 2 13-15 0.19517544 0.147561368
## 3 13-16 0.09997322 0.576122817
## 4 14-15 0.19003497 0.147561368
## 5 14-16 0.24747771 0.002768099
## 6 15-16 0.24265267 0.023063064
```

The results of the Kolmogorov-Smirnov test above seem to suggest the largemouth bass population is not stable (*Or is it? WTF do I know May be differences in sample design (i.e. different sites sampled each year repeating every 5 yrs)consider pooling year??*). The length frequency distribution **is significant different** between the years 2013 and 2014 (D = 0.24, P = 0.008), 2014 and 2016 (D = 0.25, P = 0.003), and 2015 and 2016 (D = 0.24, P = 0.023). There is **no significant differnce** between the length frequency distributions for 2013 and 2015 (D = 0.20, P = 0.128), 2013 and 2016 (D = 0.10, P = 0.576), and 2014 and 2015 (D = 0.19, P = 0.148).